Terms used **regression** and **k harmonic means** and **data mining** Found **989** of **175,083**

Sort results by: relevance
Display results: expanded form
Save results to a Binder
Search Tips
Open results in a new window
Try an Advanced Search
Try this search in The ACM Guide

Results 1 - 20 of 200 Result page: **1** 2 3 4 5 6 7 8 9 10 next
Best 200 shown Relevance scale

1 Data mining criteria for tree-based regression and classification
Andreas Buja, Yung-Seop Lee
August 2001 **Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining**
Publisher: ACM Press
Full text available: pdf(545.21 KB) Additional Information: full citation, abstract, references, citings, index terms

This paper is concerned with the construction of regression and classification trees that are more adapted to data mining applications than conventional trees. To this end, we propose new splitting criteria for growing trees. Conventional splitting criteria attempt to perform well on both sides of a split by attempting a compromise in the quality of fit between the left and the right side. By contrast, we adopt a data mining point of view by proposing criteria that search for interesting subsets ...

Keywords: Boston Housing data, CART, Pima Indians Diabetes data, splitting criteria

2 Industry/government track papers: Effective localized regression for damage detection in large complex mechanical structures
Aleksandar Lazarevic, Ramdev Kanapady, Chandrika Kamath
August 2004 **Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining KDD '04**
Publisher: ACM Press
Full text available: pdf(597.35 KB) Additional Information: full citation, abstract, references, index terms

In this paper, we propose a novel data mining technique for the efficient damage detection within the large-scale complex mechanical structures. Every mechanical structure is defined by the set of finite elements that are called structure elements. Large-scale complex structures may have extremely large number of structure elements, and predicting the failure in every single element using the original set of natural frequencies as features is exceptionally time-consuming task. Traditional data m ...

Keywords: clustering, damage detection, localized regression, mechanical structures, structure elements

3 Research papers: data mining: An integrated approach for scaling up classification and prediction algorithms for data mining
Patricia E. N. Lutu
September 2002 **Proceedings of the 2002 annual research conference of the South African institute of computer scientists and information technologists on Enablement through technology SAICSIT '02**
Publisher: South African Institute for Computer Scientists and Information Technologists
Full text available: pdf(197.71 KB) Additional Information: full citation, abstract, references, index terms

Classification and prediction algorithms for machine learning typically require all training data to be resident in memory during decision tree construction. Typically, a flat file is created from database or data warehouse data and loaded into memory for processing. This severely limits the scalability of these algorithms to practical data mining tasks. Some attempts have been made by researchers to implement disk-based algorithms which can handle much larger training sets. Both approaches suff ...

Keywords: classification, classification trees, data mining, decision tree induction, knowledge discovery in databases, machine learning, prediction, regression trees

4 Classification and regression: money *can* grow on trees

Johannes Gehrke, Wie-Yin Loh, Raghu Ramakrishnan

August 1999 **Tutorial notes of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining**

Publisher: ACM Press

Full text available: pdf(2.95 MB) Additional Information: full citation, abstract, references, citings, index terms

With over 800 million pages covering most areas of human endeavor, the World-wide Web is a fertile ground for data mining research to make a difference to the effectiveness of information search. Today, Web surfers access the Web through two dominant interfaces clicking on hyperlinks and searching via keyword queries This process is often tentative and unsatisfactory Better support is needed for expressing one's information need and dealing with a search result in more structured ways than ...

5 A survey on wavelet applications in data mining

Tao Li, Qi Li, Shenghuo Zhu, Mitsunori Ogihara

December 2002 **ACM SIGKDD Explorations Newsletter**, Volume 4 Issue 2

Publisher: ACM Press

Full text available: pdf(330.06 KB) Additional Information: full citation, abstract, references, citings

Recently there has been significant development in the use of wavelet methods in various data mining processes. However, there has been written no comprehensive survey available on the topic. The goal of this is paper to fill the void. First, the paper presents a high-level data-mining framework that reduces the overall process into smaller components. Then applications of wavelets for each component are reviewd. The paper concludes by discussing the impact of wavelets on data mining research an ...

6 Multi Relational Data Mining (MRDM): Multi-relational data mining: an introduction

Sašo Džeroski

July 2003 **ACM SIGKDD Explorations Newsletter**, Volume 5 Issue 1

Publisher: ACM Press

Full text available: pdf(1.71 MB) Additional Information: full citation, abstract, references, citings

Data mining algorithms look for patterns in data. While most existing data mining approaches look for patterns in a single data table, multi-relational data mining (MRDM) approaches look for patterns that involve multiple tables (relations) from a relational database. In recent years, the most common types of patterns and approaches considered in data mining have been extended to the multi-relational case and MRDM now encompasses multi-relational (MR) association rule discovery, MR decision tree ...

Keywords: inductive logic programming, multi-relational data mining, relational association rules, relational data mining, relational decision trees, relational distance-based methods

7 The true lift model: a novel data mining approach to response modeling in database marketing

Victor S. Y. Lo

December 2002 **ACM SIGKDD Explorations Newsletter**, Volume 4 Issue 2

Publisher: ACM Press

Full text available: pdf(119.81 KB) Additional Information: full citation, abstract, references

In database marketing, data mining has been used extensively to find the optimal customer targets so as to maximize return on investment. In particular, using marketing campaign data, models are typically developed to identify characteristics of customers who are most likely to respond. While these models are helpful in identifying the likely responders, they may be targeting customers who have decided to take the desirable action or not regardless of whether they receive the campaign contact (e ...

Keywords: customer development, customer relationship management, data mining, database marketing, interaction effect, knowledge discovery, predictive modeling, response modeling, treatment effect, true lift, upselling and cross-selling

8 Tree induction vs. logistic regression: a learning-curve analysis

Claudia Perlich, Foster Provost, Jeffrey S. Simonoff

December 2003 **The Journal of Machine Learning Research**, Volume 4

Publisher: MIT Press

Full text available: pdf(263.37 KB) Additional Information: full citation, abstract, references, citings, index terms

Tree induction and logistic regression are two standard, off-the-shelf methods for building models for classification. We present a large-scale experimental comparison of logistic regression and tree induction, assessing classification accuracy and the quality of rankings based on class-membership probabilities. We use a learning-curve analysis to examine the relationship of these measures to the size of the training set. The results of the study show several things. (1) Contrary to some prior o ...

9 Research track posters: Privacy preserving regression modelling via distributed computation

Ashish P. Sanil, Alan F. Karr, Xiaodong Lin, Jerome P. Reiter

August 2004 **Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining KDD '04**

Publisher: ACM Press

Full text available: pdf(264.04 KB) Additional Information: full citation, abstract, references, index terms

Reluctance of data owners to share their possibly confidential or proprietary data with others who own related databases is a serious impediment to conducting a mutually beneficial data mining analysis. We address the case of vertically partitioned data -- multiple data owners/agencies each possess a few attributes of every data record. We focus on the case of the agencies wanting to conduct a linear regression analysis with complete records without disclosing values of their own attributes. Thi ...

Keywords: data confidentiality, data integration, regression, secure multi-party computation

10 Constraints in data mining: *SPARTAN*: using constrained models for guaranteed-error semantic compression

Shivnath Babu, Minos Garofalakis, Rajeev Rastogi

June 2002 **ACM SIGKDD Explorations Newsletter**, Volume 4 Issue 1

Publisher: ACM Press

Full text available: pdf(259.12 KB) Additional Information: full citation, abstract, references, citings

While a variety of lossy compression schemes have been developed for certain forms of digital data (e.g., images, audio, video), the area of lossy compression techniques for arbitrary data tables has been left relatively unexplored. Nevertheless, such techniques are clearly motivated by the ever-increasing data collection rates of modern enterprises and the need for effective, guaranteed-quality approximate answers to queries over massive relational data sets.In this paper, we propose *SPARTAN* ...

11 Data Mining with optimized two-dimensional association rules

Takeshi Fukuda, Yasuhiko Morimoto, Shimichi Morishita, Takeshi Tokuyama

June 2001 **ACM Transactions on Database Systems (TODS)**, Volume 26 Issue 2

Publisher: ACM Press

Full text available: pdf(947.41 KB) Additional Information: full citation, abstract, references, index terms

We discuss data mining based on association rules for two numeric attributes and one Boolean attribute. For example, in a database of bank customers, Age and Balance are two numeric attributes, and CardLoan is a Boolean attribute. Taking the pair (Age, Balance) as a point in two-dimensional space, we consider an association rule of the form Age,Balance $\in P \Rightarrow$

Keywords: association rules, convex hull searching, data mining, image segmentation, matrix searching

12 Evolutionary algorithms in data mining: multi-objective performance modeling for direct marketing

Siddhartha Bhattacharyya

August 2000 **Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining**

Publisher: ACM Press

Full text available: pdf(115.20 KB) Additional Information: full citation, references, citings, index terms

Keywords: Pareto-optimal models, data mining, database marketing, evolutionary computation, multiple objectives

13 Statistics and data mining techniques for lifetime value modeling

D. R. Mani, James Drew, Andrew Betz, Piew Datta

August 1999 **Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining**

Publisher: ACM Press

Full text available: pdf(1.16 MB) Additional Information: full citation, references, citings, index terms

Keywords: lifetime value, neural networks, proportional hazards regression, survival analysis, tenure prediction

14 Industrial/government track: The data mining approach to automated software testing

Mark Last, Menahem Friedman, Abraham Kandel

August 2003 **Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining**

Publisher: ACM Press

Full text available: pdf(296.40 KB) Additional Information: full citation, abstract, references, index terms

In today's industry, the design of software tests is mostly based on the testers' expertise, while test automation tools are limited to execution of pre-planned tests only. Evaluation of test outputs is also associated with a considerable effort by human testers who often have imperfect knowledge of the requirements specification. Not surprisingly, this manual approach to software testing results in heavy losses to the world's economy. The costs of the so-called "catastrophic" software failures ...

Keywords: automated software testing, finite element solver, info-fuzzy networks, input-output analysis, regression testing

15 Quantifiable data mining using ratio rules

Flip Korn, Alexandros Labrinidis, Yannis Kotidis, Christos Faloutsos

February 2000 **The VLDB Journal — The International Journal on Very Large Data Bases**, Volume 8 Issue 3-4

Publisher: Springer-Verlag New York, Inc.

Full text available: pdf(451.80 KB) Additional Information: full citation, abstract, citings, index terms

Association Rule Mining algorithms operate on a data matrix (e.g., customers $\times$

products) to derive association rules [AIS93b, SA96]. We propose a new paradigm, namely, *Ratio Rules*, which are quantifiable in that we can measure the "goodness" of a set of discovered rules. We also propose the "guessing error" as a measure of the "goodness", that is, the root-mean-square error of the reconstructed values of the cells of the given matrix, when we pre ...

Keywords: Data mining, Forecasting, Guessing error, Knowledge discovery

16 Analog synthesis & design methodology: Remembrance of circuits past: macromodeling by data mining in large analog design spaces

Hongzhou Liu, Amit Singhee, Rob A. Rutenbar, L. Richard Carley
June 2002 **Proceedings of the 39th conference on Design automation**
Publisher: ACM Press

Full text available: pdf(583.79 KB) Additional Information: full citation, abstract, references, citings, index terms

The introduction of simulation-based analog synthesis tools creates a new challenge for analog modeling. These tools routinely visit 103 to 105 fully simulated circuit solution candidates. What might we do with all this circuit data? We show how to adapt recent ideas from large-scale data mining to build models that capture significant regions of this visited performance space, parameterized by variables manipulated by synthesis, trained by the data points visited during synthesis. Experimental ...

17 Statistical inference and data mining

Clark Glymour, David Madigan, Daryl Pregibon, Padhraic Smyth
November 1996 **Communications of the ACM**, Volume 39 Issue 11
Publisher: ACM Press
Full text available: pdf(383.29 KB) Additional Information: full citation, references, citings, index terms

18 Poster papers: SECRET: a scalable linear regression tree algorithm

Alin Dobra, Johannes Gehrke
July 2002 **Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining**
Publisher: ACM Press

Full text available: pdf(752.32 KB) Additional Information: full citation, abstract, references, citings, index terms

Developing regression models for large datasets that are both accurate and easy to interpret is a very important data mining problem. Regression trees with linear models in the leaves satisfy both these requirements, but thus far, no truly scalable regression tree algorithm is known. This paper proposes a novel regression tree construction algorithm (SECRET) that produces trees of high quality and scales to very large datasets. At every node, SECRET uses the EM algorithm for Gaussian mixtures to ...

19 WAPS, a data mining support environment for the planning of warranty and goodwill costs in the automobile industry

E. Hotz, G. Nakhaeizadeh, B. Petzsche, H. Spiegelberger
August 1999 **Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining**
Publisher: ACM Press
Full text available: pdf(356.65 KB) Additional Information: full citation, references, citings, index terms

Keywords: data mining, regression analysis, structural change

20 Exploring data mining implementation

Karim K. Hirji
July 2001 **Communications of the ACM**, Volume 44 Issue 7

Publisher: ACM Press

Full text available: pdf(101.20 KB) html(40.65 KB) Additional Information: full citation, references, citings, index terms, review

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2006 ACM, Inc.
Terms of Usage Privacy Policy Code of Ethics Contact Us

Useful downloads: Adobe Acrobat QuickTime Windows Media Player Real Player

Results for "((regression and k-harmonic means and data mining)<in>metadata)"
Your search matched 1 of 1340257 documents.
A maximum of 100 results are displayed, 25 to a page, sorted by Relevance in Descending order.

e-mail printer

» Search Options

View Session History

New Search

Modify Search

((regression and k-harmonic means and data mining)<in>metadata) Search

Check to search only within this results set

» Key

IEEE JNL	IEEE Journal or Magazine
IEE JNL	IEE Journal or Magazine
IEEE CNF	IEEE Conference Proceeding
IEE CNF	IEE Conference Proceeding
IEEE STD	IEEE Standard

Display Format: (•) Citation () Citation & Abstract

view selected items Select All Deselect All

1. **Regression clustering**
Zhang, B.;
Data Mining, 2003. ICDM 2003. Third IEEE International Conference on
19-22 Nov. 2003 Page(s):451 - 458
AbstractPlus | Full Text: PDF(318 KB) IEEE CNF
Rights and Permissions

Indexed by Inspec

© Copyright 2006 IEEE – All Rights

Google Scholar BETA

regression and "k-harmonic means" and "data | Search | Advanced Scholar Search | Scholar Preferences | Scholar Help

The "AND" operator is unnecessary -- we include all search terms by default. [details]

Scholar Results **1 - 10** of about **13** for **regression and "k-harmonic means" and "data mining"**. (**0.11** seconds)

Categorizing Visitors Dynamically by Fast and Robust Clustering of Access Logs - group of 4 »
All articles Recent articles
V Estivill-Castro, J Yang - Lecture Notes in Computer Science, 2001 - Springer
... So Web-site owners discover more knowledge by **data mining** techniques, and im- plement them for Web-site analysis settings [19]. ...
Cited by 10 - Web Search

Fast and Robust General Purpose Clustering Algorithms - group of 3 »
VS Estivill-Castro, JS Yang - **Data Mining** and Knowledge Discovery, 2004 - Springer
... Efficient clustering is a fundamental task in **data mining** where the goal is to discover patterns with large data sets (thousands or millions of records) that ...
Cited by 15 - Web Search - BL Direct

Learning Simple Relations: Theory and Applications - group of 6 »
P Berkhin, JD Becher - ... the The Second SIAM International Conference on **Data Mining**, 2002 - siam.org
... Keywords – Clustering, K-Means, Mutual Information, Web Mining, **Data Mining** ... The introduced model is well fitted for different **data mining** applications. ...
Cited by 13 - View as HTML - Web Search

Algorithms for clustering high dimensional and distributed data - group of 3 »
T Li - Intelligent Data Analysis, 2003 - IOS Press
... among attributes in categorical ("basket") databases [3], called association mining, is one of the most fundamental and most popular problems in **data mining** ...
Cited by 6 - Web Search

Multi-agent systems and distributed **data mining** - group of 2 »
C Giannella, R Bhargava, H Kargupta - Cooperative Information Agents VIII: 8th International ..., 2004 - Springer
... the one presented in [3], but extend it to **K-harmonic means**. ... Multi-agent Systems and Distributed **Data Mining** 7 ... to pro- duce a global **regression**-based classifier ...
Cited by 1 - Web Search - BL Direct

[BOOK] Temporal, Spatial, and Spatio-Temporal **Data Mining**: First International Workshop, Tsdm 2000, Lyon, ...
JF Roddick, K Hornsby - 2001 - books.google.com
... Hsu, and Umesh Dayal presented a center- based clustering algorithm, the **K-Harmonic Means**. ... a paper by Richard Povinelli on time series **data mining** as applied ...
Cited by 1 - Web Search

Regression clustering - group of 3 »
B Zhang - **Data Mining**, 2003. ICDM 2003. Third IEEE International ..., 2003 - ieeexplore.ieee.org
... q (>=1) will put the **regression**'s error ... is more general than the **K-Harmonic Means** clustering algorithm ... IEEE International Conference on **Data Mining** (ICDM'03 ...
Web Search

Directed Research Project PerfExplorer: Parallel Performance Analysis Using **Data Mining** Techniques
KA Huck - cs.uoregon.edu
... tried three methods, cost minimization, **regression** models and ... The primary **data mining** operation we are performing in the ... One such method is **k-harmonic-means**. ...
View as HTML - Web Search

Distributed **Data Mining** and Agents
JC da Silva, C Giannella, R Bhargava, H Kargupta, ... - dfki.de

... that have appeared in the **data mining** literature ... the one presented in [10], but extend it to **K -harmonic means**. ... to pro- duce a global **regression**-based classifier ...
View as HTML - Web Search

Fast and Robust General Purpose Clustering Algorithms
J YANG - **Data Mining** and Knowledge Discovery, 2004 - kluweronline.com
... Efficient clustering is a fundamental task in **data mining** where the goal is to discover patterns with large data sets (thousands or millions of records) that ...
Web Search

©2006 Google